

Demo Day Final Report

Machine Learning with TensorFlow Training Professional Academy Digital Talent Scholarship 2022

Group Number	BMKG 1
Dataset	BMKG/Weather
Name - DTS ID	<ol style="list-style-type: none">1. Adam Aji Nugroho - 152236035100-712. Gradytama Elkana - 152236035100-10373. Rindang Muharza Viawan - 152236035101-3904. Aldhi Prambudi - 152236035100-1103

Selected Theme: Weather

Title of the Project: Prediksi Curah Hujan untuk Prakiraan Musim (studi kasus: Kota Denpasar)

Background Summary:

Informasi cuaca dan iklim bermanfaat bagi publik dan sektor-sektor lainnya seperti pertanian, penerbangan, konstruksi, industri, kelautan, pariwisata dan lain-lain. Sektor pertanian misalnya, membutuhkan informasi musim yakni kapan datangnya musim hujan dan kemarau sehingga membantu untuk menentukan jenis tanaman.

Pada riset ini akan dilakukan prediksi curah hujan sebagai dasar prakiraan musim. Prediksi curah hujan yang dilakukan akan per dasarian atau 10 harian dan juga bulanan. Dari prediksi tersebut dapat dilihat apakah curah hujan berada diatas suatu ambang batas. Jika berada diatas ambang batas maka dianggap musim hujan sedangkan jika dibawah maka musim kemarau, namun juga melihat nilai sebelum dan sesudahnya. Misalkan jika terdapat nilai dibawah ambang batas diantara nilai diatas ambang batas maka dianggap masih musim hujan. Ambang batas yang dipakai adalah curah hujan 50 mm / dasarian.

Dataset yang digunakan tidak hanya satu melainkan dua dataset. Dataset dari Tim DTS dan BMKG. Dataset tersebut dipilih karena berada pada zona musim yang sama. Data pada zona musim yang sama dapat digunakan secara bersama untuk prakiraan musim. Dataset tersebut berisi data 29 tahunan Kota Denpasar dan Kab. Badung. Data tersebut akan diolah menjadi data perhari untuk mengurangi kompleksitas model. Data preprocessing, eda dan cleansing juga dilakukan sesuai dengan karakter dari datasetnya.

Desain model menggunakan Sequence Model dengan multi input/multi output. Output berupa curah hujan satu tahun kedepan dalam harian yang nantinya akan dilakukan agregasi per bulan dan dasarian. Pada saat prediksi digunakan data perhari karena jika sudah di agregasi diawal menjadi perbulan / dasarian maka data point akan menjadi sangat sedikit sedangkan deep learning model membutuhkan data yang banyak berbeda dengan klasikal model. Secara mendetail model menerima 365 titik timestep dan prediksi 365 titik kedepan dimana titik merepresentasikan hari. Hasil riset menunjukkan bidirectional GRU memberikan hasil yang optimum untuk permasalahan yang ada, selain itu jumlah unit juga diperhatikan agar tidak bottleneck dengan output.

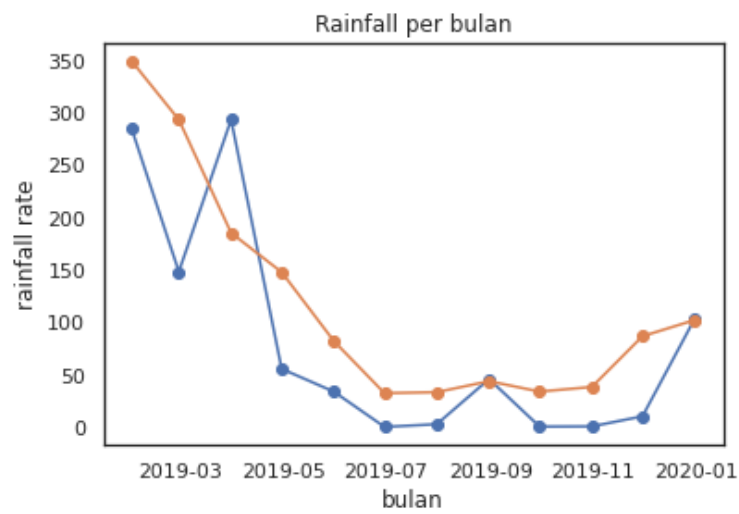
Tema ini dipilih karena kemanfaatannya dan masih minimnya penggunaan ML / DL dalam riset prakiraan cuaca dan iklim di Indonesia dan dunia. Oleh karena itu, harapan kami riset ini menjadi satu diantara pembuka bagi riset-riset lainnya.

Please specifically mention what you've done:

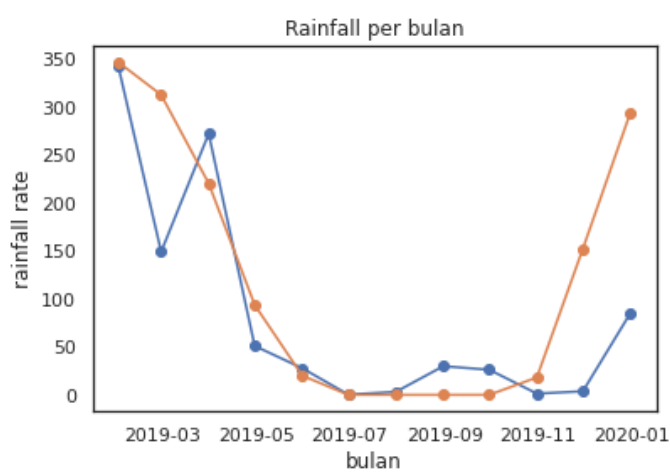
1. **Adam Aji Nugroho - 152236035100-71** : PIC EDA dataset 1, PIC Dataset preprocessing dataset 1, eksplorasi model dan evaluasi dataset 1, koordinasi tim, download dataset 2
2. **Gradytama Elkana - 152236035100-1037** : PIC Desain model dataset 2, eksplorasi model, PIC Testing dataset 2, menyiapkan dokumen proposal dan Analisis akhir,
3. **Rindang Muharza Viawan - 152236035101-390** : PIC EDA dataset 2, PIC Dataset preprocessing dataset 2, eksplorasi model dan evaluasi dataset 2
4. **Aldhi Prambudi - 152236035100-1103** : PIC Desain model dataset 1, PIC Testing dataset 1, menyiapkan dokumen proposal, final report & Github

Screenshot Output Model :

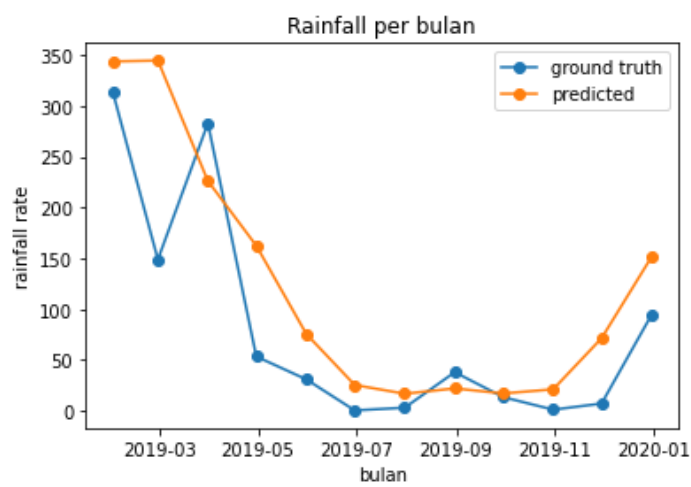
- Model untuk dataset 1



- Model untuk dataset 2



- Rata-rata model dataset 1 dan 2



Github Repo Link: https://github.com/AldhiPrambudi/BMKG_1/

10-Min Video Presentation Link:
https://github.com/AldhiPrambudi/BMKG_1/blob/main/Project%20Report/BMKG_1.mp4

Quotes : *There is no such thing as bad weather, only different kinds of good weather.*

Project Reference:

Dataset :

1. Dataset Pengamatan Meteorologi Daerah Ngurah Rai

(<https://dataonline.bmkg.go.id/home>)

Dataset dapat diunduh dari link dataonline.bmkg.go.id, namun untuk menghindari potensi pelanggaran penyebaran dataset maka dataset yang telah kami unduh diletakkan pada link berikut dengan password yang tertera.

<https://drive.google.com/file/d/1R-jBrox89cn9A9JI3ZF7-3Vd4obN89qe/view?usp=sharing>

Pass RAR : BMKG1@DTSPROA

2. Dataset Openweather Denpasar

(<https://drive.google.com/drive/folders/12fKqOJMKc7SJXJvEoHCgP4b7q9u0bED1>)

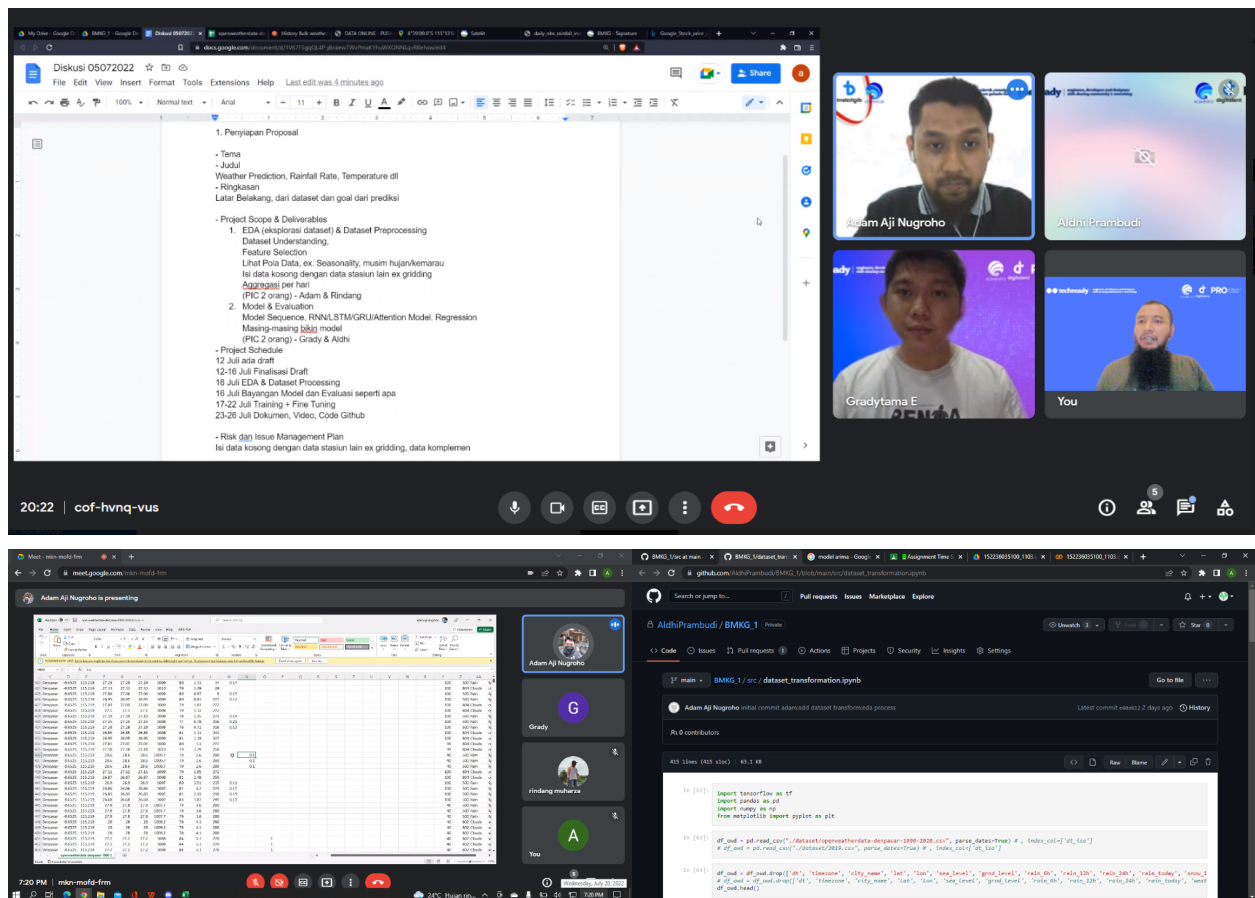
Artikel :

- Daftar Istilah Klimatologi (<http://balai3.denpasar.bmkg.go.id/daftar-istilah-musim>)
- Buletin Prakiraan Musim Hujan 2020/2021 Prov. Bali - 2020 - BMKG (<http://iklim.bali.bmkg.go.id/wp-content/uploads/2020/10/PMH-2021.pdf>)

Paper Terkait:

- D.S. Roy, "Forecasting The Air Temperature at a Weather Station Using Deep Neural Networks", 9th International Young Scientist Conference on Computational Science, 2020
- Chhetri M, Kumar S, Pratim Roy P, Kim B-G. Deep BLSTM-GRU Model for Monthly Rainfall Prediction: A Case Study of Simtokha, Bhutan. *Remote Sensing*. 2020; 12(19):3174. <https://doi.org/10.3390/rs12193174>

Documentation:



The screenshot shows a Google Meet session with four participants: Adam Aji Nugroho, Aldhi Pambudi, Gradytama E, and You. The main window displays a Google Docs document titled "Diskusi 05072022" with a table of contents and a list of topics. The right sidebar shows four participants: Adam Aji Nugroho, Aldhi Pambudi, Gradytama E, and You. The bottom window shows a Google Meet interface with a list of participants and a code editor displaying Python code for dataset transformation.

Table of Contents:

- 1. Penyajian Proposal
 - Tema
 - Judul
 - Weather Prediction, Rainfall Rate, Temperature dll
 - Ringkasan
 - Latar Belakang, dari dataset dan goal dari prediksi
- Project Scope & Deliverables
 - 1. EDA (eksplorasi dataset) & Dataset Preprocessing
 - Dataset Understanding
 - Feature Selection
 - Lihat Pola Data, ex. Seasonality, musim hujan/kemarau
 - Isi data kosong dengan data stasiun lain ex gridding
 - Aggragasi per hari
 - (PIC 2 orang) - Adam & Rindang
 - 2. Model & Evaluation
 - Model Sequence, RNN/LSTM/GRU/Attention Model, Regression
 - Masing-masing bikin model
 - (PIC 2 orang) - Grady & Aldhi
- Project Schedule
 - 12 Juli ada draft
 - 12-16 Juli Finalisasi Draft
 - 16 Juli EDA & Dataset Processing
 - 16 Juli Bayangan Model dan Evaluasi seperti apa
 - 17-22 Juli Training + Fine Tuning
 - 23-26 Juli Dokumen, Video, Code Github
- Risk dan Issue Management Plan
- Isi data kosong dengan data stasiun lain ex gridding, data komplemen

Python Code:

```

import tensorflow as tf
import pandas as pd
import numpy as np
from tensorflow.keras.layers import LSTM

df_train = pd.read_csv('dataset/openweatherdata-denpasar-2020-2021.csv', parse_dates=True)
df_test = pd.read_csv('dataset/2020-2021.csv', parse_dates=True)

df_train = df_train.drop(['lat', 'lon', 'sea_level', 'grnd_level', 'rain_1h', 'rain_3h', 'rain_6h', 'rain_12h', 'rain_24h', 'rain_48h', 'rain_72h', 'rain_96h', 'rain_120h', 'rain_144h', 'rain_168h', 'rain_192h', 'rain_216h', 'rain_240h', 'rain_264h', 'rain_288h', 'rain_312h', 'rain_336h', 'rain_360h', 'rain_384h', 'rain_408h', 'rain_432h', 'rain_456h', 'rain_480h', 'rain_504h', 'rain_528h', 'rain_552h', 'rain_576h', 'rain_600h', 'rain_624h', 'rain_648h', 'rain_672h', 'rain_696h', 'rain_720h', 'rain_744h', 'rain_768h', 'rain_792h', 'rain_816h', 'rain_840h', 'rain_864h', 'rain_888h', 'rain_912h', 'rain_936h', 'rain_960h', 'rain_984h', 'rain_1008h', 'rain_1032h', 'rain_1056h', 'rain_1080h', 'rain_1104h', 'rain_1128h', 'rain_1152h', 'rain_1176h', 'rain_1200h', 'rain_1224h', 'rain_1248h', 'rain_1272h', 'rain_1296h', 'rain_1320h', 'rain_1344h', 'rain_1368h', 'rain_1392h', 'rain_1416h', 'rain_1440h', 'rain_1464h', 'rain_1488h', 'rain_1512h', 'rain_1536h', 'rain_1560h', 'rain_1584h', 'rain_1608h', 'rain_1632h', 'rain_1656h', 'rain_1680h', 'rain_1704h', 'rain_1728h', 'rain_1752h', 'rain_1776h', 'rain_1800h', 'rain_1824h', 'rain_1848h', 'rain_1872h', 'rain_1896h', 'rain_1920h', 'rain_1944h', 'rain_1968h', 'rain_1992h', 'rain_2016h', 'rain_2040h', 'rain_2064h', 'rain_2088h', 'rain_2112h', 'rain_2136h', 'rain_2160h', 'rain_2184h', 'rain_2208h', 'rain_2232h', 'rain_2256h', 'rain_2280h', 'rain_2304h', 'rain_2328h', 'rain_2352h', 'rain_2376h', 'rain_2400h', 'rain_2424h', 'rain_2448h', 'rain_2472h', 'rain_2496h', 'rain_2520h', 'rain_2544h', 'rain_2568h', 'rain_2592h', 'rain_2616h', 'rain_2640h', 'rain_2664h', 'rain_2688h', 'rain_2712h', 'rain_2736h', 'rain_2760h', 'rain_2784h', 'rain_2808h', 'rain_2832h', 'rain_2856h', 'rain_2880h', 'rain_2904h', 'rain_2928h', 'rain_2952h', 'rain_2976h', 'rain_3000h', 'rain_3024h', 'rain_3048h', 'rain_3072h', 'rain_3096h', 'rain_3120h', 'rain_3144h', 'rain_3168h', 'rain_3192h', 'rain_3216h', 'rain_3240h', 'rain_3264h', 'rain_3288h', 'rain_3312h', 'rain_3336h', 'rain_3360h', 'rain_3384h', 'rain_3408h', 'rain_3432h', 'rain_3456h', 'rain_3480h', 'rain_3504h', 'rain_3528h', 'rain_3552h', 'rain_3576h', 'rain_3600h', 'rain_3624h', 'rain_3648h', 'rain_3672h', 'rain_3696h', 'rain_3720h', 'rain_3744h', 'rain_3768h', 'rain_3792h', 'rain_3816h', 'rain_3840h', 'rain_3864h', 'rain_3888h', 'rain_3912h', 'rain_3936h', 'rain_3960h', 'rain_3984h', 'rain_4008h', 'rain_4032h', 'rain_4056h', 'rain_4080h', 'rain_4104h', 'rain_4128h', 'rain_4152h', 'rain_4176h', 'rain_4200h', 'rain_4224h', 'rain_4248h', 'rain_4272h', 'rain_4296h', 'rain_4320h', 'rain_4344h', 'rain_4368h', 'rain_4392h', 'rain_4416h', 'rain_4440h', 'rain_4464h', 'rain_4488h', 'rain_4512h', 'rain_4536h', 'rain_4560h', 'rain_4584h', 'rain_4608h', 'rain_4632h', 'rain_4656h', 'rain_4680h', 'rain_4704h', 'rain_4728h', 'rain_4752h', 'rain_4776h', 'rain_4800h', 'rain_4824h', 'rain_4848h', 'rain_4872h', 'rain_4896h', 'rain_4920h', 'rain_4944h', 'rain_4968h', 'rain_4992h', 'rain_5016h', 'rain_5040h', 'rain_5064h', 'rain_5088h', 'rain_5112h', 'rain_5136h', 'rain_5160h', 'rain_5184h', 'rain_5208h', 'rain_5232h', 'rain_5256h', 'rain_5280h', 'rain_5304h', 'rain_5328h', 'rain_5352h', 'rain_5376h', 'rain_5400h', 'rain_5424h', 'rain_5448h', 'rain_5472h', 'rain_5496h', 'rain_5520h', 'rain_5544h', 'rain_5568h', 'rain_5592h', 'rain_5616h', 'rain_5640h', 'rain_5664h', 'rain_5688h', 'rain_5712h', 'rain_5736h', 'rain_5760h', 'rain_5784h', 'rain_5808h', 'rain_5832h', 'rain_5856h', 'rain_5880h', 'rain_5904h', 'rain_5928h', 'rain_5952h', 'rain_5976h', 'rain_6000h, 'rain_6024h', 'rain_6048h', 'rain_6072h', 'rain_6096h', 'rain_6120h', 'rain_6144h', 'rain_6168h', 'rain_6192h', 'rain_6216h', 'rain_6240h', 'rain_6264h', 'rain_6288h', 'rain_6312h', 'rain_6336h', 'rain_6360h', 'rain_6384h', 'rain_6408h', 'rain_6432h', 'rain_6456h', 'rain_6480h', 'rain_6504h', 'rain_6528h', 'rain_6552h', 'rain_6576h', 'rain_6600h', 'rain_6624h', 'rain_6648h', 'rain_6672h', 'rain_6696h', 'rain_6720h', 'rain_6744h', 'rain_6768h', 'rain_6792h', 'rain_6816h', 'rain_6840h', 'rain_6864h', 'rain_6888h, 'rain_6912h', 'rain_6936h', 'rain_6960h', 'rain_6984h', 'rain_7008h', 'rain_7032h', 'rain_7056h', 'rain_7080h', 'rain_7104h', 'rain_7128h', 'rain_7152h', 'rain_7176h', 'rain_7200h', 'rain_7224h', 'rain_7248h', 'rain_7272h', 'rain_7296h', 'rain_7320h', 'rain_7344h', 'rain_7368h', 'rain_7392h', 'rain_7416h', 'rain_7440h', 'rain_7464h', 'rain_7488h', 'rain_7512h', 'rain_7536h', 'rain_7560h', 'rain_7584h', 'rain_7608h', 'rain_7632h', 'rain_7656h', 'rain_7680h', 'rain_7704h', 'rain_7728h', 'rain_7752h', 'rain_7776h', 'rain_7800h', 'rain_7824h', 'rain_7848h', 'rain_7872h', 'rain_7896h', 'rain_7920h', 'rain_7944h', 'rain_7968h', 'rain_7992h', 'rain_8016h', 'rain_8040h', 'rain_8064h', 'rain_8088h, 'rain_8112h', 'rain_8136h', 'rain_8160h', 'rain_8184h', 'rain_8208h', 'rain_8232h', 'rain_8256h', 'rain_8280h', 'rain_8304h', 'rain_8328h', 'rain_8352h', 'rain_8376h', 'rain_8400h', 'rain_8424h', 'rain_8448h', 'rain_8472h', 'rain_8496h', 'rain_8520h', 'rain_8544h', 'rain_8568h', 'rain_8592h', 'rain_8616h', 'rain_8640h', 'rain_8664h', 'rain_8688h, 'rain_8712h', 'rain_8736h', 'rain_8760h', 'rain_8784h', 'rain_8808h', 'rain_8832h', 'rain_8856h', 'rain_8880h', 'rain_8904h', 'rain_8928h', 'rain_8952h', 'rain_8976h', 'rain_9000h', 'rain_9024h', 'rain_9048h', 'rain_9072h', 'rain_9096h', 'rain_9120h', 'rain_9144h', 'rain_9168h', 'rain_9192h', 'rain_9216h', 'rain_9240h', 'rain_9264h', 'rain_9288h, 'rain_9312h', 'rain_9336h', 'rain_9360h', 'rain_9384h', 'rain_9408h', 'rain_9432h', 'rain_9456h', 'rain_9480h', 'rain_9504h', 'rain_9528h', 'rain_9552h', 'rain_9576h', 'rain_9600h', 'rain_9624h', 'rain_9648h', 'rain_9672h', 'rain_9696h', 'rain_9720h', 'rain_9744h', 'rain_9768h', 'rain_9792h', 'rain_9816h', 'rain_9840h', 'rain_9864h', 'rain_9888h, 'rain_9912h', 'rain_9936h', 'rain_9960h', 'rain_9984h', 'rain_10000h']
df_test = df_test.drop(['lat', 'lon', 'sea_level', 'grnd_level', 'rain_1h', 'rain_3h', 'rain_6h', 'rain_12h', 'rain_24h', 'rain_48h', 'rain_72h', 'rain_96h', 'rain_120h', 'rain_144h', 'rain_168h', 'rain_192h', 'rain_216h', 'rain_240h', 'rain_264h', 'rain_288h', 'rain_312h', 'rain_336h', 'rain_360h', 'rain_384h', 'rain_408h', 'rain_432h', 'rain_456h', 'rain_480h', 'rain_504h', 'rain_528h', 'rain_552h', 'rain_576h', 'rain_600h', 'rain_624h', 'rain_648h', 'rain_672h', 'rain_696h', 'rain_720h', 'rain_744h', 'rain_768h', 'rain_792h', 'rain_816h', 'rain_840h', 'rain_864h', 'rain_888h, 'rain_912h', 'rain_936h', 'rain_960h', 'rain_984h', 'rain_10000h']

```

